

## Fast optimization schemes for feature selection in analog circuits fault diagnosis

**Abstract.** The aim of this work is to systematize the knowledge resulting from research on the impact of the feature selection on the quality of diagnostic procedures in the diagnosis of nonlinear systems. Particular attention was devoted to the selection of appropriate comparative criteria and optimization algorithms necessary for the selection of defects in the studied nonlinear systems, so that the inclusion of the elements in the process of detection and location of single and multiple catastrophic failures is possible to the highest degree. Basing on the research and simulations results, the fast, "low-costs" method for feature selection using new data quality indexes was invented and tested on real circuits examples.

**Streszczenie.** Celem pracy jest usystematyzowanie wiedzy wynikającej z badań realnego wpływu selekcji cech na jakość procedur wykrywania uszkodzeń w diagnostyce układów nieliniowych. Szczególna uwaga została poświęcona na dobór właściwych kryteriów porównawczych i algorytmów optymalizacyjnych niezbędnych w procesie wyboru atrybutów uszkodzeń badanych układów nieliniowych tak, aby w jak największym stopniu możliwe było uwzględnienie tolerancji elementów w procesie detekcji i lokalizacji jednokrotnych i wielokrotnych uszkodzeń katastroficznych. Opierając się na wynikach analiz i symulacji opracowano i przetestowano na przykładach, szybki w działaniu algorytm selekcji cech wykorzystujący nowe indeksy oceny jakości zbioru danych. (**Prosta optymalizacja procesu selekcji cech w diagnostyce układów analogowych**).

**Keywords:** feature selection, BPSO algorithm, tolerance, classifiers.

**Keywords:** selekcja cech, binarny algorytm PSO, tolerancja, klasyfikatory.

doi:10.12915/pe.2014.07.12

### Introduction

The consideration of the elements' tolerances in the diagnostics of non-linear circuits is still a current and not yet satisfactorily solved problem. Deviations of the studied circuits elements' can be taken into account by either the experimental or analytical determination of tolerance areas. Another approach is the use of classifiers that include the range of elements' values at the stage of procedures for damage recognition learning. This kind of approach requires big sets of data. What follows is the need to optimize the algorithms in the direction of minimizing the duration of the preparatory procedures, mainly by reducing the size of the input data. Such optimization can be provided by modern methods of feature selection (measuring points) of the diagnostic systems based on both - the data mining techniques and efficient heuristic optimization techniques.

The aim of this work is to systematize the knowledge resulting from research on the impact of the feature selection on the quality of diagnostic procedures in the diagnosis of nonlinear systems. Particular attention was devoted to the selection of appropriate comparative criteria and optimization algorithms necessary for the selection of defects in the studied nonlinear systems, so that the inclusion of the elements' tolerances in the process of detection and location of single and multiple catastrophic failures is possible to the highest degree.

### Evaluation criteria for the data quality factor for the set of features

Optimization of sets of measurement points is one of the most important problems of diagnostics of electronic systems. It is the subject of many scientific papers, including i.e. [3],[21],[22]. However, relatively little is known about the selection of a set of attributes for diagnosis, which takes into account the tolerances of the system components. The most reliable methods here seem to be based on testing data sets for specific classifiers. They are extremely time-consuming procedures. In this paper, we devote ourselves to the analysis of the simpler methods, largely based on the structure of the data collected. This approach ensures that the optimized set of data will improve the efficiency of a wide range of classifiers, starting with those based on the ranking lists, through a huge family

of classifiers using neural networks to the heuristic methods based on evolutionary techniques.

The problem of assessing the quality of the results of feature selection is similar to the problem of assessing the quality of clustering results as part of the pattern recognition algorithms. Quality indicators of data clustering are generally divided into external (supervised), relative (relative) and internal (unsupervised) [9]. The latter group of quality index is best suited for use in the evaluation of the required data sets of attributes for the comparison of measurement in the diagnosis of electronic systems. Many well-known algorithms for clustering features (eg.: agglomeration) are used to measure the degree of similarity or differentiation. Appropriate measures to define the degree of similarity or differences significantly affect the accuracy of the data quality assessment. Most definitions of such measures can be found in the work of Sneath [20], Anderbrga [1], Everitt [2]. A good indicator of the quality of clustering, based on the information from the interior of the analyzed data set, may be, so called, silhouette index (SI), defined by P.J Rousseeuw and successfully implemented in the works for the classification of the company's clients [16]. Use of the SI indicator is presented, among others, in works of L.F. Largo-Fernandez and F.Corbacho [7] and M.Ming-Tsao and B.Mirkin [10]. Originally, clustering quality indicators are used to optimize the number of classes (groups). In the version used in this article, indicators may allow the evaluation of the selection of the number and type of attributes (features) of the externally imposed structure of the classes.

To the group of the undoubtedly most famous clustering quality indices belongs also the Fisher index [19],[23]. This index, together with SI and invented GRA based index [14] will be briefly define and describe in the next subsections.

### FISHER Index

Among all the statistical criteria enabling us to give feature rankings, Fishers's criterion seems to be the most known and popular. For a given feature  $X_i$  containing  $K$  classes, let us denote the set of instances in class  $k$  as  $X_i^k$  and number of instances per class as  $n_k$ ,  $k = 1, 2, \dots, c$  [24]. The Fisher score (F-score, F-index) of individual feature is defined as

$$(1) \quad F_i = \frac{S_B^i}{S_W^i}$$

where

$$(2) \quad S_B^i = \sum_{k=1}^C n_k (\bar{X}_i^k - \bar{X}_i)^2,$$

$$S_W^i = \sum_{k=1}^C \sum_{x \in X_i^k} (x - \bar{X}_i)^2,$$

and  $\bar{X}_i^k, \bar{X}_i$  are the mean values of the i-th feature in  $X_i^k$  and  $X_i$  respectively. Overall index, enabling us to define class separability of a feature set, can be estimated from the relation

$$(3) \quad \hat{F} = \text{trace}(S_W^{-1} S_B)$$

where  $S_W$  represents the within-scatter matrix and  $S_B$  represents the between-scatter matrix of the given data set  $X$  [23]. Measure (3) serves as a very good criterion for features selection, but, unfortunately, it is computationally expensive. Hence, in many practical problems, the similar concept, called Sebestyen criterion [19], is applied. Using additive property of the coefficients (2) in terms of particular feature, overall modified *cluster validity index* (defining how informative the data set is) may be expressed in the following simplified forms

$$(4) \quad F_{SEB1} = \frac{1}{n} \sum_{k=1}^{nf} (S_B^i - a S_W^i) q_k,$$

$$F_{SEB2} = \frac{\sum_{k=1}^{nf} (S_B^i) q_k}{\sum_{k=1}^{nf} (S_W^i) q_k}, \quad \sum_{k=1}^{nf} q_k = n$$

where  $q_k$  is equal to 1, when we choose k-th feature, and equal to 0 - if otherwise,  $nf$  denotes number of all available features, subscript *cof* indicates subset of features to be estimated.

#### SILHOUETTE Index

Formal definition for this data quality factor is as follows [16]. Let

$$\mathbf{X} = \{\mathbf{X}_1, \dots, \mathbf{X}_{NI}\}$$

is the data set (NI- number of all instances)

$$C = \{C_1, \dots, C_K\}$$

denotes its clustering into K clusters and

$$C_j = \{\mathbf{X}_1^j, \dots, \mathbf{X}_{m_j}^j\}$$

define j-th cluster ( $j=1,2,\dots,K$ ,  $m_j$  cluster's numerical amount, number of the vectors in the cluster). Finally, the silhouette of the class  $C_j$  is given by the following formula:

$$(5) \quad S_j = \frac{1}{m_j} \sum_{i=1}^{m_j} s_i^j,$$

where

$$(6) \quad s_j = \frac{b_i^j - a_i^j}{m_j}$$

Average distance between i-th vector from  $C_j$  and the other vectors from the same cluster, and the minimum distance between i-th vector from  $C_j$  and all the vectors belonging to the clusters  $C_k$ ,  $k=1,\dots,K$ ,  $k \neq j$  are given by the following expressions respectively:

$$(7) \quad a_i^j = \frac{1}{m_j - 1} \sum_{\substack{k=1 \\ k \neq j}}^{m_j} d(\mathbf{X}_i^j, \mathbf{X}_k^j),$$

$$b_i^j = \min_{\substack{n=1,\dots,K \\ n \neq j}} \left\{ \frac{1}{m_n} \sum_{\substack{k=1 \\ k \neq j}}^{m_j} d(\mathbf{X}_i^j, \mathbf{X}_k^j) \right\}, i=1,2,\dots,m_j$$

Global Silhouette index has the form:

$$(8) \quad F_{SIL} = \frac{1}{K} \sum_{j=1}^K S_j,$$

#### GRA-based indices

##### Filter approach

To measure discriminatory power of the chosen set of features we may also use the concept presented in [12], [13] and based on relational degree of the data set [4], [8]. The invented feature set quality factor can be expressed in the following form:

$$(9) \quad F_G(\mathbf{a}) = \text{mean}_{i=1,\dots,M} \{\Delta_i(\mathbf{a})\} + P(\beta)$$

where

$$(10) \quad \Delta_i(\mathbf{a}) = \left| 1 - \max_{\substack{j=1,\dots,M \\ j \neq i}} \Gamma_{ij}^0 \right|, i=1,\dots,M,$$

and

$$(11) \quad \beta = \sum_i^M (\Delta_i(\mathbf{a}) \leq \varepsilon)$$

Penalty function  $P(\beta)$  depends on the number of classes with small ( $\leq \varepsilon$ ) separability factor (10). The value of  $P(\beta)$  is zero when  $\beta$  is zero and increases rapidly for  $\beta > 2$ .

Binary vector

$$(12) \quad \mathbf{a} = [a_1 \ a_2 \ \dots \ a_M]^T$$

indicates the combination of chosen features and auxiliary matrix  $\Gamma^0(\mathbf{a})$  [12] where each j-th column ( $j=1,2,\dots,M$ ) represents the relational degree (13) of circuit under test in terms of the reference set of features (testing points) for nominal values of the defined state used as testing fault [12], [13].

$$(13) \quad \Gamma_{ij}^0 = \sum_{k=1}^N \alpha_k (\gamma_i)_{jk}$$

where N is a number of features defined for the circuit under test, M – number of all investigated states of the

$$(14) \quad (\gamma_i)_{jk} = \frac{\mu_o + \delta \eta_o}{|(x_i)_o^k - x_j^k| + \beta \eta_o}, \delta \in (0,1),$$

$$(15) \quad \mu_o = \min_j \min_k |(x_i)_o^k - x_j^k|, \eta_o = \max_j \max_k |(x_i)_o^k - x_j^k|$$

where

$$x_j^k \in X, k = 1, \dots, N, j = 1, \dots, M$$

represents circuit under test measurements and

$$(x_i)_o^k = x_i^k, k = 1, 2, \dots, N$$

defines reference set of features for auxiliary  $\Gamma^0(\mathbf{a})$  [12].  $N$  is a number of features (attributes) defined for the circuit under test,  $M$  – number of all investigated states of the circuit, including nominal one.

#### Wrapper approach

The filter type feature selection described above does not need large data sets, but its ability to improve effectiveness of GRA-based ranking list classifiers is limited. Taking into account reduced data set following from small number of randomly chosen simulations, assuming that probability of each state of circuit under test is the same, the fitness function for feature selection optimization scheme may be defined as

$$(16) \quad F_{GW}(\mathbf{a}) = \frac{N_P(\mathbf{a})}{N_S * N_{CS}}$$

where:  $N_S$  - number of all circuit states,  $N_{CS}$  - number of circuit simulations (per state),  $N_P(\mathbf{a})$  - number of positive fault classifications.

#### Data quality factors comparison

To illustrate a real affect of feature selection “low-cost” methods on the behavior of exemplary ranking list classifier, the benchmark circuit from Fig.1 was investigated (elements values and types:  $R_1=47k\Omega$ ,  $R_2=430\Omega$ ,  $R_3=47k\Omega$ ,  $R_4=2.2k\Omega$ ,  $R_5=100k\Omega$ ,  $R_6=680\Omega$ ,  $R_7=47k\Omega$ ,  $V_{CC}=15V$ , two bipolar transistors BC109). Catastrophic faults of all the resistive elements and state with nominal values of the elements were assumed ( $R_S=10E-6\Omega$  for short circuit,  $R_O=10E+12\Omega$  for open one). The initial set of features contains measurements of five independent node voltages and two voltage sources currents for three different values of supplying voltage source. All the data sets were prepared with use of SPICE simulator (ICAP-4, Winspice3). For different tolerances of elements and different number of Monte Carlo simulations per circuit state, the following quality factors were tested: both Sebestyn modified coefficients (4), two forms of SI data quality factor (8) and invented GRA based index (9-11). Table 1 contains random probe of the best three measures investigated features’ sets (for 8 features from all of 21). The best correlation with effectiveness of GRA classifier [14] does Silhouette Index (without preliminary normalization of data) posses (Fig.2). However, also the GRA based index (Fig.3) enables us to estimate behavior of the classification procedure.

Hence, the main procedure of feature selection algorithm should be based on some version of Silhouette measure. However, in the case when number of available measurement points (features) is large, selection based on data quality factors for all the features requires very time-consuming numerical procedures (see Table 2). Therefore, the ranking procedure using measure  $F_G$  (9), can be applied as a preliminary feature set reduction method. The balance sheet of time-consumption for various ranking procedures for single feature (for different numbers of circuit’s states and Monte Carlo trials) states that GRA is about 10 times

faster than SI method applied for 15 classes (states) and 100 samples per state (see Table II).

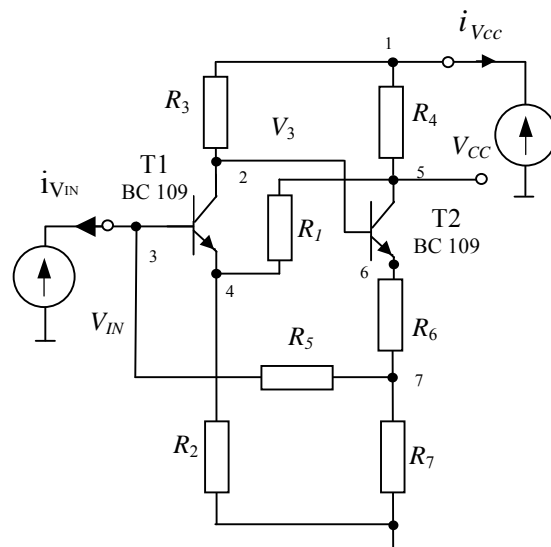


Fig.1. Benchmark Darlington circuit [12]

Table 1. Comparison of feature selection validity indices for test circuit of Fig.1 (exemplary part of simulations’ results set)

Features set	$F_{SEB2}$	$F_{SIL}$	$F_G$	$F_{GW}$
49	0,028621	0,915343	0,886187	0,992481
50	0,046826	0,812141	0,565754	0,842105
:	:	:	:	:
60	0,154648	0,936872	0,670522	0,984962
61	0,923340	0,651136	0,340099	0,857143
62	0,250848	0,904400	0,521651	0,969925
63	0,154648	0,936872	0,670522	0,984962
:	:	:	:	:
113	0,428204	0,917339	0,558471	0,887218
114	0,405076	0,840932	0,502062	0,857143
:	:	:	:	:

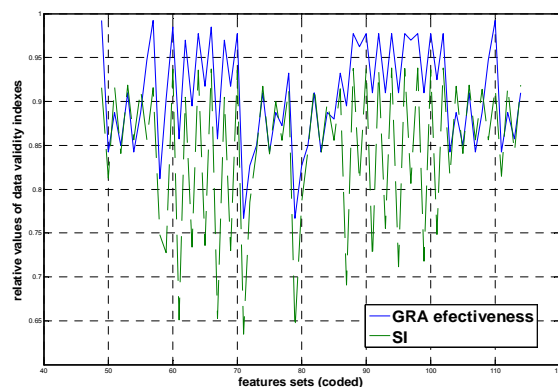


Fig.2. Correlation between effectiveness of the ranking list classifier and Silhouette Index

Table 2. Time consumption of the procedures calculating chosen data quality indices for test circuit of Fig.1 (single feature measure)

$F_{SEB2}$	$F_{SIL}$	$F_G$	$F_{GMC}$
47.20	38.81ms	0.33ms	0.38ms

#### Optimization algorithms

Taking into account conclusions following from results of tests and simulations, two “low-cost” multilevel optimization algorithms are proposed and verified. The first one is

dedicated to improve effectiveness of fault diagnosis classifiers based on ranking lists.

Algorithm 1

**STEP\_0:** Define measurement points (features), circuit faults (classes)

**STEP\_1:** Perform circuit simulation for nominal values of elements.

**STEP\_2:** Calculate  $F_G$  factor (9) for all available features. Cancel 20-25% the worst features.

**STEP\_3:** Choose number of Monte Carlo trials ( $NMC \leq 10$ ) and perform circuit simulations to prepare Testing Set.

For  $i=1:N$

**STEP\_4:** Start binary optimization algorithm with fitness function (9)-(11) (population size:  $50 \times \text{number of variables}$ ).

**STEP\_5:** Store the population containing the best individual

end

**STEP\_6:** Rank the stored  $N$  populations with use of formula (16) and Testing Set

If we replace fitness function from **STEP\_4** with the function based on (16) applied to another testing set of simulating data (called Validating Set,  $NMC \leq 10$ ), optimization process of feature selection will be more time consuming but also more effective (Algorithm 2).

Comparing a few binary optimization techniques we decided to use a discrete binary version of the Particle Swarm Optimization (DPSO) algorithm originally designed and introduced by J.Kennedy and R.C.Eberhart [6]. The application of DPSO was proposed i.e. in [15]. Each potential solution in DPSO, called a particle, is assigned a randomized velocity and differ from real-valuated version in the definition of the particles values, which must be restricted to the one from the set  $\{0,1\}$ . The particles and their velocities are updates as follows:

$$(17) \quad v_{id}^{t+1} = \omega v_{id}^t + r_1^t (P_{best,id}^t - x_{id}^t) + r_2^t (G_{best,id}^t - x_{id}^t)$$

$$x_{id}^{t+1} = \begin{cases} 0, & \text{for } rand \geq \text{sigmoid}(v_{id}^{t+1}) \\ 1, & \text{for others} \end{cases}$$

$$\text{sigmoid}(v_{id}^{t+1}) = \left[ 1 + \exp(-v_{id}^{t+1}) \right]^{-1}$$

where:  $x_{id}^t$  and  $v_{id}^t$  are respectively the particle and the velocity in the  $t$ -th iteration,  $P_{best,id}^t$  and  $G_{best,id}^t$  are the best positions that respectively the particle and the hole group of particles ever had up to  $t$ -th iteration,  $r_1^t$  and  $r_2^t$  are the random values in  $t$ -th iteration,  $rand$  is random number selected from a uniform distribution in  $[0,1]$ ,  $\omega$  is the coefficient, that provides a balance between local and global exploration.

In case when simulations of the circuit for different set of parameters varying within intervals defined by the assumed values of tolerances are not available, measures based on Monte Carlo analysis may be replaced by simplified factor (16), which uses only measurements for nominal values of circuit parameters only. Hence, the simplest and fastest version of feature selection method should be presented in the following form:

Algorithm 3

**STEP\_0:** Define measurement points and circuit faults.

**STEP\_1:** Perform circuit simulation for nominal values of elements.

**STEP\_2:** Calculate  $F_G$  factor (9) for all available features. Cancel 20-25% the worst features.

For  $i=1:N$

**STEP\_3:** Start binary optimization algorithm with fitness function (9)-(11) (population size:  $50 \times \text{number of variables}$ ).

**STEP\_5:** Store the population containing the best individual

end

**STEP\_6:** Rank the stored  $N$  populations with use of formula (9)-(11).

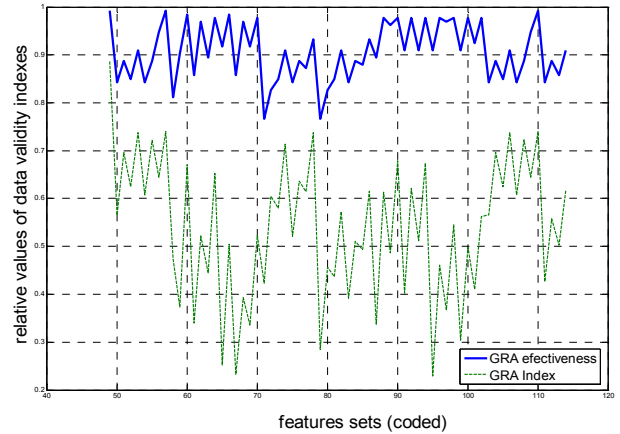


Fig.3. Correlation between effectiveness sets of the ranking list classifier and Silhouette Index.

### Numerical simulations

The results of many simulations of different analog bipolar and CMOS circuits are very similar to the results obtained for the benchmark circuit from Fig.1. Table III contains comparison of fault recognition error rate calculated for the set of features achieved from three described in the previous section versions of optimization scheme. Last column contains the average error rate for randomly chosen 50 sets of eight features (from all available 21). Performing the first variant of the optimization method (see Section 4) the locally selected optimal set of features contains five test points:

$$\mathfrak{S}_1 = \{1,2,3,4,10,19\}$$

From the Algorithm 2 the following features were obtained:

$$\mathfrak{S}_2 = \{1,3,4,12,17,19\},$$

whereas the set

$$\mathfrak{S}_3 = \{2,3,4,10,11,12\}$$

follows from the third version of the method (features 7,9,14,16 and 21 were removed from the initial set after preliminary ranking stage, feature codes are defined in Table 5). Error ratio was simulated by use of another data set (containing at least 4 times more samples than Testing Set) classified by GRA-based ranking list algorithm [14].

Table 3. Effectiveness of Ranking List Classifier for different set of features

Algorithm 1	Algorithm 2	Algorithm 3	Random selection (average ratio)
8%	5%	11%	24%

Let us limit the results of comparative analysis of the feature selection algorithms working with neural network supervised classifiers to the RBFN (Radial Basis Function Network) standard method [11]. To preserve clear test conditions the same number of features and learning epochs of RBF network (375) were ensured. Moreover, no additional optimizations of the RBF-network were done. Table 4 contains error rates for optimal sets of features obtained by the use of Algorithm 1 and 2 in comparison to average value the rate for 10 randomly selected feature sets (each 8 features).

Table 4. RBF-based classifier Error Rate

	FS(Alg.2)	FS(Alg.1)	Random Selection set
Error rate	19,47%	21,23%	37,33%
MSE	2.4e-6	3.5e-6	6.3e-4

Table 5. Coding scheme for features

$V_{IN}$	V(2)	V(4)	V(5)	V(6)	V(7)	$I_{VCC}$	$I_{VIN}$
2.2V	1	2	3	4	6	6	7
4V	8	9	10	11	12	13	14
5V	15	16	17	18	19	21	21

All the computer simulations were performed with use of Matlab (2013a)-WinSpice3 environment.

## Conclusion

The aim of this work was to systematize the knowledge resulting from research on the impact of the "low-cost" feature selection methods on the quality of diagnostic procedures. Particular attention was devoted to the selection of appropriate comparative criteria. Simulations and numerical experiments lead to the conclusion, that Silhouette Index, considered one of the best data quality measure, may be successfully replaced by GRA-based factors (Chapter *GRA-based indices*). Application of these indices as fitness function for modified binary PSO optimization algorithm, improves effectiveness of ranking list classifiers as well as different Neural Network classifiers with supervised learning. Moreover, invented and discussed here GRA-based factor (9-11) enables us to select features (testing points) even when data sets containing results of circuit's simulations with parameters varying around nominal values (tolerance included) are not available.

## Acknowledgments

This work was supported by the statutory activities of Technical University of Lodz 112/1/DzS/2014.

## REFERENCES

- [1] Anderberg M.R., Cluster analysis for applications, Academic Press, New York, San Francisco, London, 1973.
- [2] Everitt B.S., Cluster Analysis, Edward Arnold, London, 1993.
- [3] Golonek T., Rutkowski J., Genetic-algorithm-based method for optimal analog test points selection, *IEEE Trans. Cir. Syst. II*, vol.54, pp.117-121, 2007.
- [4] Julong D., Introduction of Grey System, *Journal of Grey System*, vol.1, no.1, 1989.
- [5] Kennedy J., Eberhard R.C., Shi J.H., Swarm Intelligence. Academic Press, 2001.

- [6] Kennedy J., Eberhard R.C., A discrete binary version of the particle swarm algorithm, *IEEE Conference on Systems, Man, and Cybernetics*, 1997.
- [7] Largo-Fernandez L.F., Corbacho F., Normality-based validation for crisp clustering, *Pattern Recognition*, 43, 3, 782-795, 2010.
- [8] Liu S., Forrest J.Y., Advances in Grey Systems Research, Springer-Verlag Berlin Heidelberg, 2010.
- [9] Migdał-Najman K., Ocena jakości wyników grupowania przegląd bibliografii, *Przegląd statystyczny*, R. LVIII, Zeszyt 3-4, 2011.
- [10] Ming-Tso Chiang M., Mirkin B. Intelligent choice of the number of clusters in k-means clustering: an experimental study with different cluster spreads, *Journal of classification*, 27, 1, 3-40, 2010.
- [11] Ossowski S., Sieci neuronowe do przetwarzania informacji, OWPW, Warszawa 2006.
- [12] Ossowski M., Korzybski M., Improving analog circuit fault diagnosis by testing points selection, *Proceedings of International Conf. On Signals and Electronic Systems, ICSES 2012*, IEEE Xplore Digital Library
- [13] Ossowski M., Improving neural network methods for time domain fault analysis of nonlinear analog circuits by feature selection, *Elektronika 2010*, R.51 nr 12 pp.36-39, 2010.
- [14] Ossowski M., Korzybski M., Data mining based algorithm for analog circuits fault diagnosis, *Przegląd Elektrotechniczny*, R. 89 NR 2a/2013, 285-287.
- [15] Perez R.E., Behdinan K., Particle swarm approach for structural design optimization. *Computers and Structures*, vol.85, pp.1579-1588, 2007.
- [16] Petrovic S., A comparison between the Silhouette Index and The Davies-Bouldin Index in labeling IDC Clusters, *Proceedings of the 11th NORDEC*, Linköping, 2006.
- [17] Prasad V.C and Babu N.S.C., Selection of test nodes for analog fault diagnosis in dictionary approach, *IEEE Trans. Inst. Measurement*, vol.49, 1289-1297, 2000.
- [18] Rousseeuw P.J., Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, *Journal of computational and applied mathematics*, 20, 1, 53-65, 1987.
- [19] Sobczak W., Malina W., Methods of selection and reduction of information, WNT, Warsaw, 1985.
- [20] Sneath P.H.A., Sokal R.R., Numerical Taxonomy. The principles and practice of numerical classification, W.H. Freeman & Company, San Francisco, CA, 1973.
- [21] Starzyk J.A., Liu D., Liu Z.H., Nelson D.E., and Rutkowski J., Entropy-based optimum test points selection for analog fault dictionary techniques, *IEEE Trans. Inst. Measurement*, vol.53, pp.754-761, 2004.
- [22] Tadeusiewicz M., Hałgas S., Korzybski M., Multiple catastrophic fault diagnosis of analog circuits considering the component tolerances, *International Journal on Circuit Theory and Application*,
- [23] Wang S., Cheng-Lin L., Zheng L., Feature selection combining Fisher Criterion and principal feature analysis, *Proc. of VI-th Int. Conf. on Machine Learning and Cybernetics*, Hong Kong, pp.1149-1154, 2007.
- [24] Yin L., Ge Y., Xiao K., Wang X., Quan X., Feature selection for high-dimensional data, *Neurocomputing 105*, pp.3-11, 2013.

**Authors:** dr inż. Marek Ossowski, Politechnika Łódzka, Instytut Systemów Inżynierii Elektrycznej, ul. Stefanowskiego 18/22, 90-924 Łódź, E-mail [marek.ossowski@p.lodz.pl](mailto:marek.ossowski@p.lodz.pl), dr inż. Marek Korzybski, Politechnika Łódzka, Instytut Systemów Inżynierii Elektrycznej, ul. Stefanowskiego 18/22, 90-924 Łódź, [marek.korzybski@p.lodz.pl](mailto:marek.korzybski@p.lodz.pl)